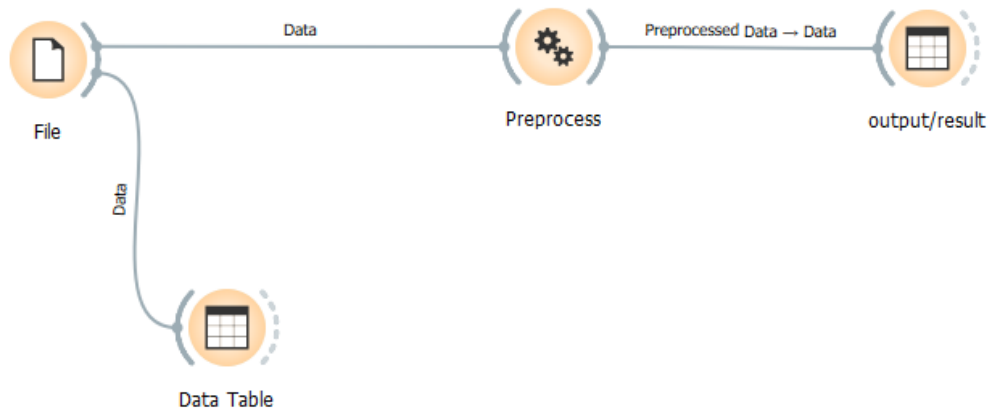


Overview of data preprocess using orange:



Dataset used: heart_disease.tab which has got the missing value of 0.2 %.

File - Orange

Source: ☒ File: heart_disease.tab ☐ URL:

Info

Heart Disease dataset
Data on the presence of heart disease in patients.

303 instance(s)
13 feature(s) (0.2% missing values)
Classification; categorical class with 2 values (no missing values)
0 meta attribute(s)

Columns (Double click to edit)

	Name	Type	Role	Values
1	age	N numeric	feature	
2	gender	C categorical	feature	female, male
3	chest pain	C categorical	feature	asymptomatic, atypical ang, non-anginal, typical ang
4	rest SBP	N numeric	feature	
5	cholesterol	N numeric	feature	
6	fasting blood sugar > 120	C categorical	feature	0, 1

Reset Apply

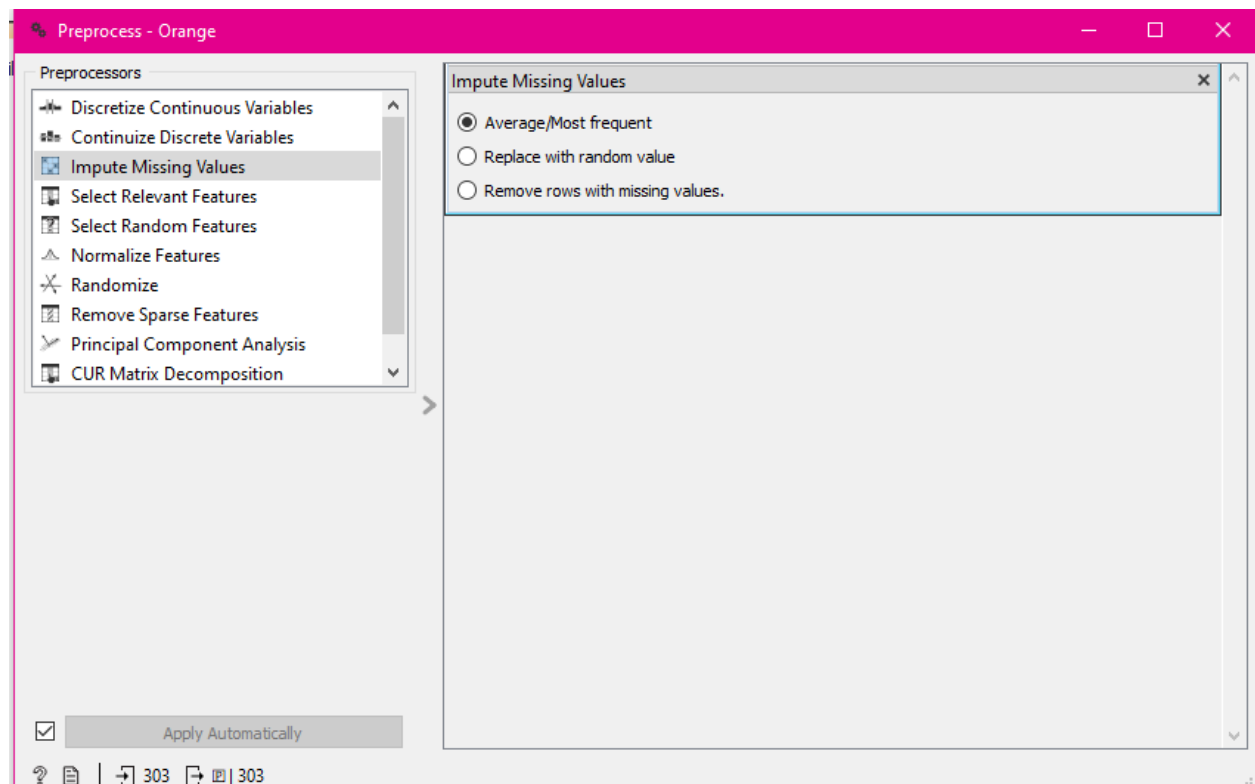
Browse documentation datasets

303

I have taken an example of a missing value which is present at serial number 167 as shown below.

	age	gender	chest pain	rest SBP	cholesterol	ng blood sugar >	rest ECG	max HR	exerc ind ang	ST by exercise	lope peak exc 5'	ajor vessels color	thal
147	57	male	asymptomatic	165	289	1	left vent hypert...	124	0		1.0	flat	3 reversible defec
148	41	male	non-anginal	112	250	0	normal	179	0		0.0	upsloping	0 normal
149	45	male	atypical ang	128	308	0	left vent hypert...	170	0		0.0	upsloping	0 normal
150	60	female	non-anginal	102	318	0	normal	160	0		0.0	upsloping	1 normal
151	52	male	typical ang	152	298	1	normal	178	0		1.2	flat	0 reversible defec
152	42	female	asymptomatic	102	265	0	left vent hypert...	122	0		0.6	flat	0 normal
153	67	female	non-anginal	115	564	0	left vent hypert...	160	0		1.6	flat	0 reversible defec
154	55	male	asymptomatic	160	289	0	left vent hypert...	145	1		0.8	flat	1 reversible defec
155	64	male	asymptomatic	120	246	0	left vent hypert...	96	1		2.2	downsloping	1 normal
156	70	male	asymptomatic	130	322	0	left vent hypert...	109	0		2.4	flat	3 normal
157	51	male	asymptomatic	140	299	0	normal	173	1		1.6	upsloping	0 reversible defec
158	58	male	asymptomatic	125	300	0	left vent hypert...	171	0		0.0	upsloping	2 reversible defec
159	60	male	asymptomatic	140	293	0	left vent hypert...	170	0		1.2	flat	2 reversible defec
160	68	male	non-anginal	118	277	0	normal	151	0		1.0	upsloping	1 reversible defec
161	46	male	atypical ang	101	197	1	normal	156	0		0.0	upsloping	0 reversible defec
162	77	male	asymptomatic	125	304	0	left vent hypert...	162	1		0.0	upsloping	3 normal
163	54	female	non-anginal	110	214	0	normal	158	0		1.6	flat	0 normal
164	58	female	asymptomatic	100	248	0	left vent hypert...	122	0		1.0	flat	0 normal
165	48	male	non-anginal	124	255	1	normal	175	0		0.0	upsloping	2 normal
166	57	male	asymptomatic	132	207	0	normal	168	1		0.0	upsloping	0 reversible defec
167	52	male	non-anginal	138	223	0	normal	169	0		0.0	upsloping	? normal
168	54	female	atypical ang	132	288	1	left vent hypert...	159	1		0.0	upsloping	1 normal
169	35	male	asymptomatic	126	282	0	left vent hypert...	156	1		0.0	upsloping	0 reversible defec
170	45	female	atypical ang	112	160	0	normal	138	0		0.0	flat	0 normal
171	70	male	non-anginal	160	269	0	normal	112	1		2.9	flat	1 reversible defec
172	53	male	asymptomatic	142	226	0	left vent hypert...	111	1		0.0	upsloping	0 reversible defec
173	59	female	asymptomatic	174	249	0	normal	143	1		0.0	flat	0 normal
174	62	female	asymptomatic	140	394	0	left vent hypert...	157	0		1.2	flat	0 normal
175	64	male	asymptomatic	145	212	0	left vent hypert...	132	0		2.0	flat	2 fixed defect
176	57	male	asymptomatic	152	274	0	normal	88	1		1.2	flat	1 reversible defec

Now in order to clean this data I used data preprocess method where this value is imputed with Average/Most frequent



After data preprocessing, I have used another table which is the result/output with no missing values.

	age	gender	chest pain	rest SBP	cholesterol	ng blood sugar >	rest ECG	max HR	exerc ind ang	ST by exercise	lope peak exc SI	for vessels color	thal
150	60	female	non-anginal	102	318	0	normal	160	0	0.0	upsloping	1	normal
151	52	male	typical ang	152	298	1	normal	178	0	1.2	flat	0	reversible defect
152	42	female	asymptomatic	102	265	0	left vent hypert...	122	0	0.6	flat	0	normal
153	67	female	non-anginal	115	564	0	left vent hypert...	160	0	1.6	flat	0	reversible defect
154	55	male	asymptomatic	160	289	0	left vent hypert...	145	1	0.8	flat	1	reversible defect
155	64	male	asymptomatic	120	246	0	left vent hypert...	96	1	2.2	downsloping	1	normal
156	70	male	asymptomatic	130	322	0	left vent hypert...	109	0	2.4	flat	3	normal
157	51	male	asymptomatic	140	299	0	normal	173	1	1.6	upsloping	0	reversible defect
158	58	male	asymptomatic	125	300	0	left vent hypert...	171	0	0.0	upsloping	2	reversible defect
159	60	male	asymptomatic	140	293	0	left vent hypert...	170	0	1.2	flat	2	reversible defect
160	68	male	non-anginal	118	277	0	normal	151	0	1.0	upsloping	1	reversible defect
161	46	male	atypical ang	101	197	1	normal	156	0	0.0	upsloping	0	reversible defect
162	77	male	asymptomatic	125	304	0	left vent hypert...	162	1	0.0	upsloping	3	normal
163	54	female	non-anginal	110	214	0	normal	158	0	1.6	flat	0	normal
164	58	female	asymptomatic	100	248	0	left vent hypert...	122	0	1.0	flat	0	normal
165	48	male	non-anginal	124	255	1	normal	175	0	0.0	upsloping	2	normal
166	57	male	asymptomatic	132	207	0	normal	168	1	0.0	upsloping	0	reversible defect
167	52	male	non-anginal	138	223	0	normal	169	0	0.0	upsloping	0.67	normal
168	54	female	atypical ang	132	288	1	left vent hypert...	159	1	0.0	upsloping	1	normal
169	35	male	asymptomatic	126	282	0	left vent hypert...	156	1	0.0	upsloping	0	reversible defect
170	45	female	atypical ang	112	160	0	normal	138	0	0.0	flat	0	normal
171	70	male	non-anginal	160	269	0	normal	112	1	2.9	flat	1	reversible defect
172	53	male	asymptomatic	142	226	0	left vent hypert...	111	1	0.0	upsloping	0	reversible defect
173	59	female	asymptomatic	174	249	0	normal	143	1	0.0	flat	0	normal
174	62	female	asymptomatic	140	394	0	left vent hypert...	157	0	1.2	flat	0	normal
175	64	male	asymptomatic	145	212	0	left vent hypert...	132	0	2.0	flat	2	fixed defect
176	57	male	asymptomatic	152	274	0	normal	88	1	1.2	flat	1	reversible defect
177	52	male	asymptomatic	108	233	1	normal	147	0	0.1	upsloping	3	reversible defect
178	56	male	asymptomatic	132	184	0	left vent hypert...	105	1	2.1	flat	1	fixed defect
179	43	male	non-anginal	130	315	0	normal	162	0	1.9	upsloping	1	normal
180	53	male	non-anginal	130	246	1	left vent hypert...	173	0	0.0	upsloping	3	normal